# Automatically Evaluating the Efficiency of Search-Based Test Data Generation

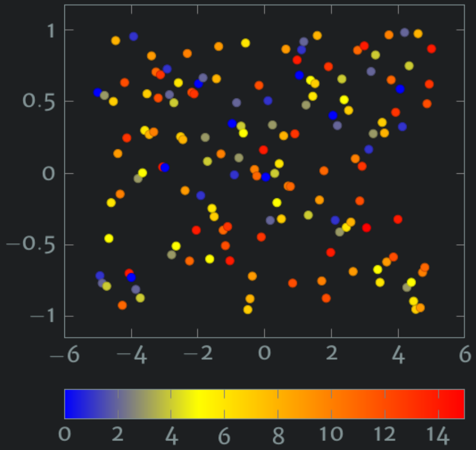
(for Relational Database Schemas)

Cody Kinneer

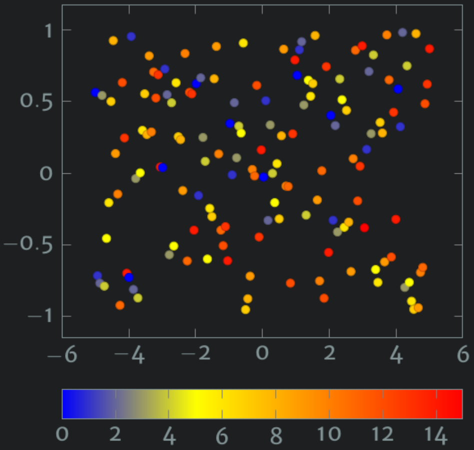
SEKE 2015

July 7, 2015

# Random Testing

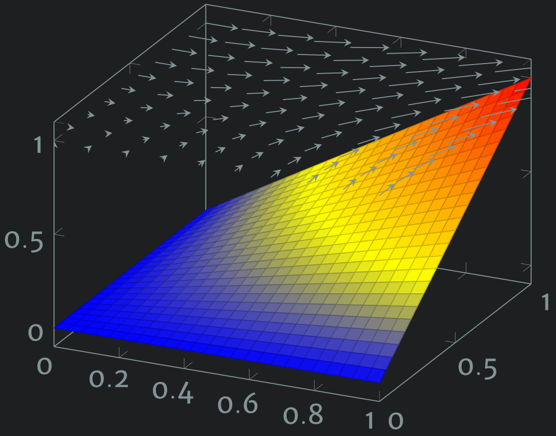


# Random Testing

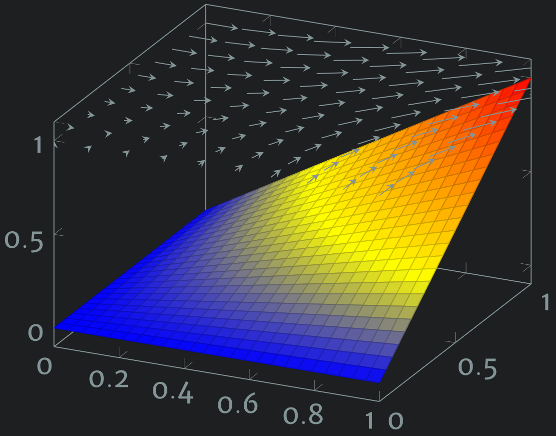



Easy to implement — and yet not always very effective!

# Search-Based Testing



# Search-Based Testing



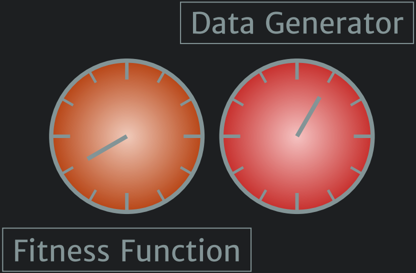

Often much more effective than random testing

# Performance of SBST

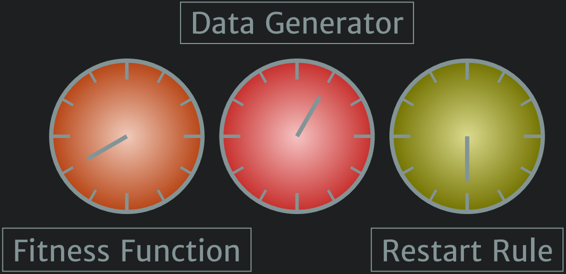
# Performance of SBST

Fitness Function

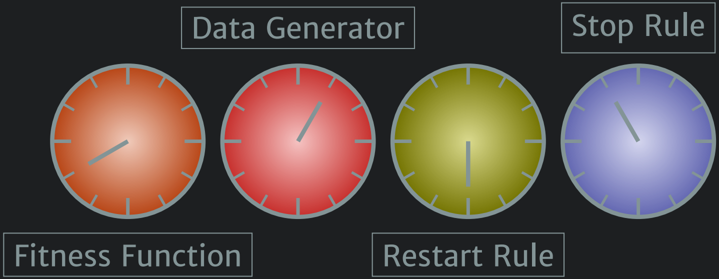
# Performance of SBST



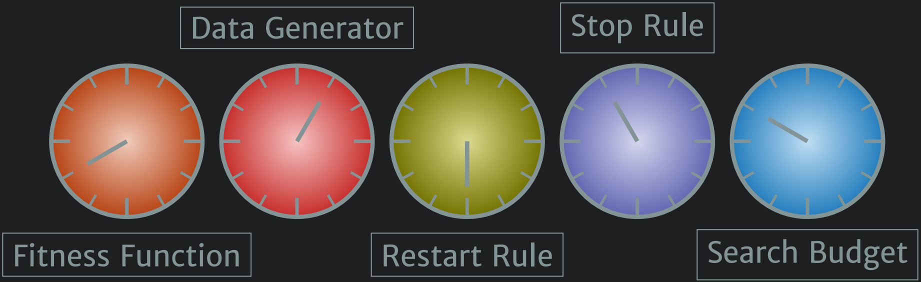
# Performance of SBST



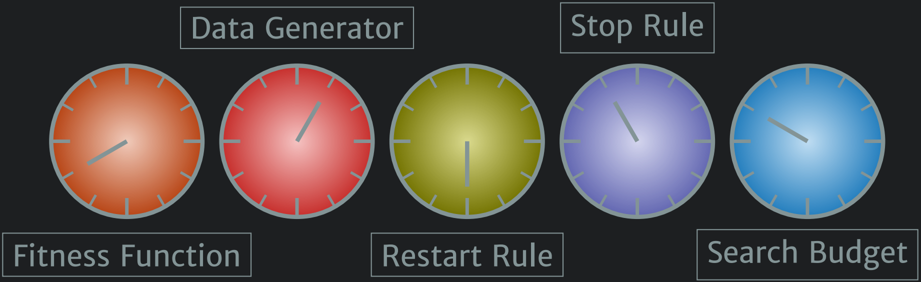
# Performance of SBST



# Performance of SBST



# Performance of SBST




How do parameter values influence the efficiency of SBST?

# Performance of SBST

$$O(\quad)$$

# Performance of SBST

$$O(?)$$

# Performance of SBST

O(?)

Analytical

# Performance of SBST

$$O(?)$$

✗Analytical

# Performance of SBST

$O(?)$

✗Analytical Empirical

# Performance of SBST

$$O(?)$$

✗Analytical ✓Empirical

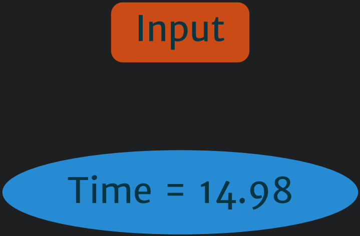
# Doubling Experiment

Input

# Doubling Experiment

Input

# Doubling Experiment



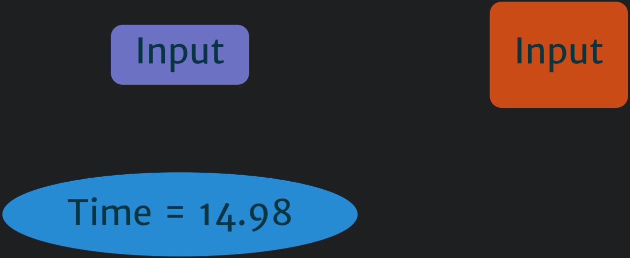
# Doubling Experiment

Input

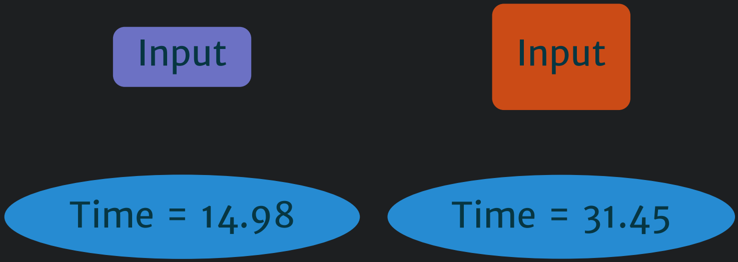
Input

Time = 14.98

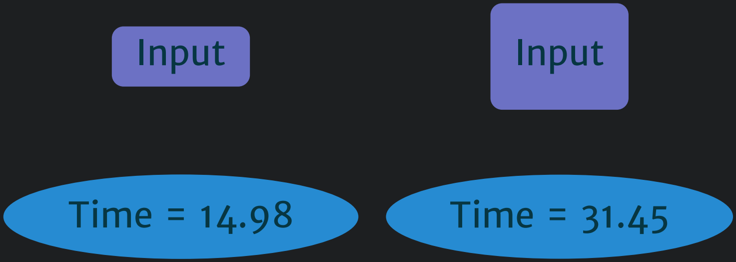
# Doubling Experiment



# Doubling Experiment



# Doubling Experiment



# Doubling Experiment

Input

Input

Time = 14.98

Time = 31.45

Ratio ≈ 2

# Doubling Experiment

Input

Input

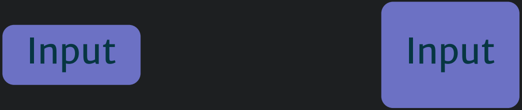
Time = 14.98

Time = 31.45

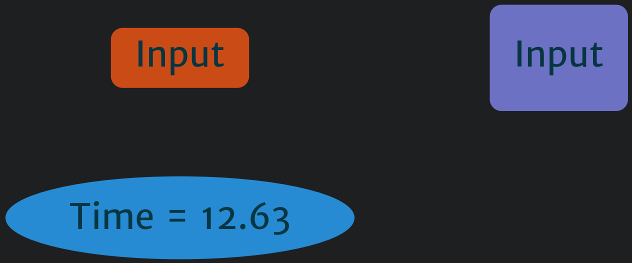
Ratio ≈ 2

Linear — O(n)

# Doubling Experiment

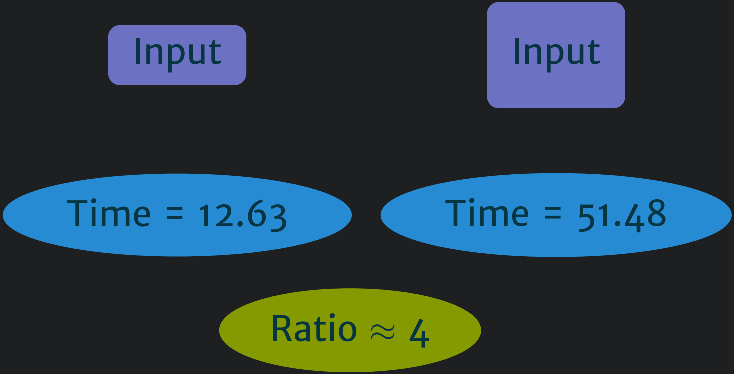


# Doubling Experiment



# Doubling Experiment

# Doubling Experiment



# Doubling Experiment

Input

Input

Time = 12.63

Time = 51.48

Ratio $\approx 4$

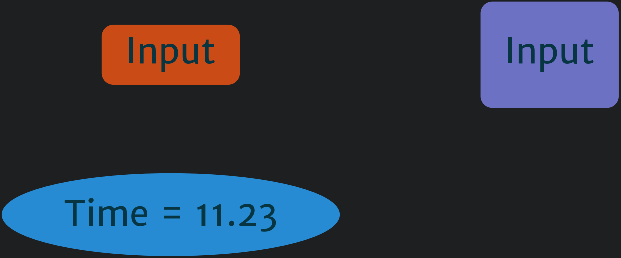
Quadratic — $O(n^2)$

# Doubling Experiment

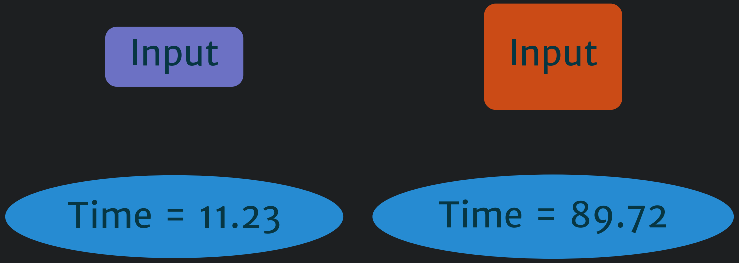
Input

Input

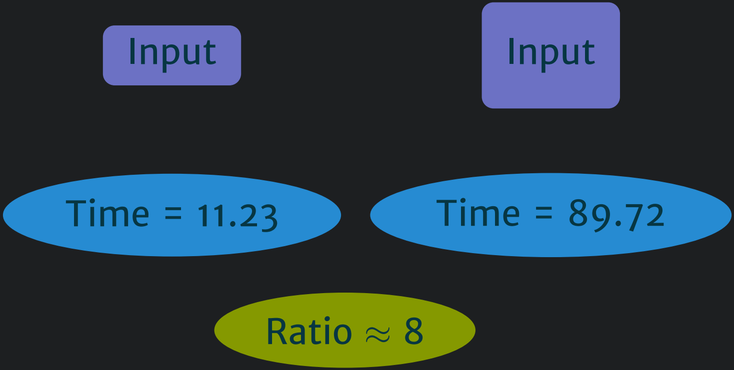
# Doubling Experiment



# Doubling Experiment



# Doubling Experiment



# Doubling Experiment

Input

Input

Time = 11.23

Time = 89.72

Ratio $\approx$ 8

Cubic — $O(n^3)$

# Relational Databases

Deployment Locations for Databases

# Relational Databases

Deployment Locations for Databases

Database Application Server

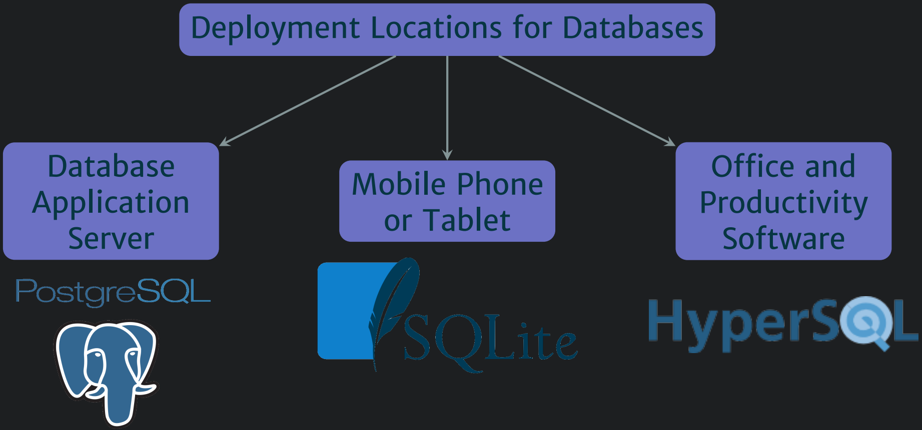
PostgreSQL

# Relational Databases

Deployment Locations for Databases

- Database Application Server
  - PostgreSQL
- Mobile Phone or Tablet
  - SQLite

# Relational Databases



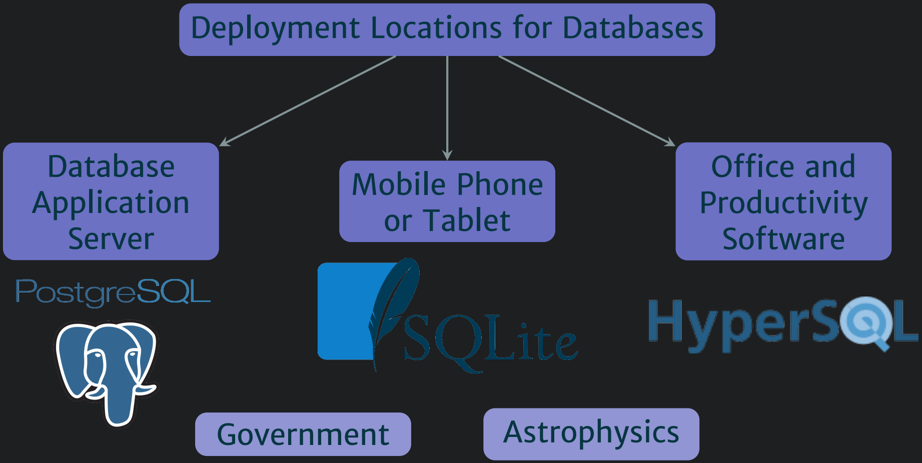
# Relational Databases

Deployment Locations for Databases

- Database Application Server
- Mobile Phone or Tablet
- Office and Productivity Software

Government

# Relational Databases



# Database Schemas

PostgreSQL

Relational Database
Management System

# Database Schemas


E-commerce

PostgreSQL

Relational Database Management System

# Database Schemas

PostgreSQL

E-commerce

Relational Database Management System

Schema

# Database Schemas

E-commerce

PostgreSQL

Relational Database Management System

Schema

State

# Database Schemas

PostgreSQL

E-commerce

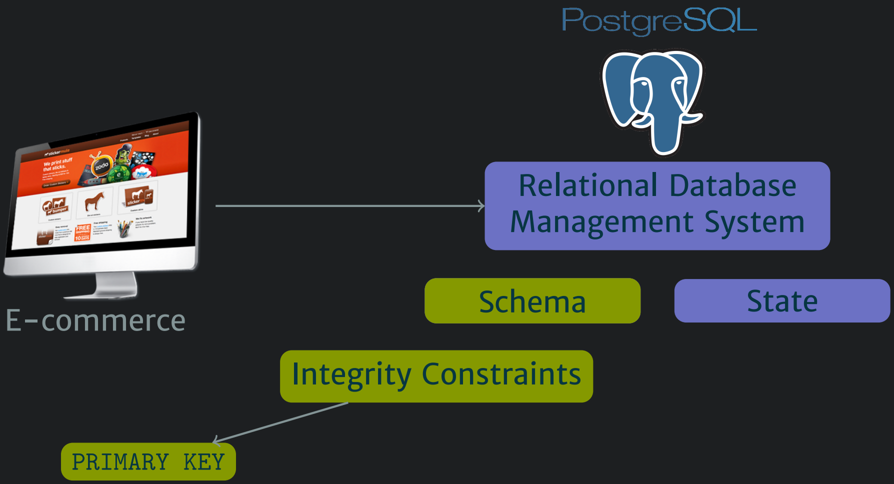
Relational Database Management System

Schema

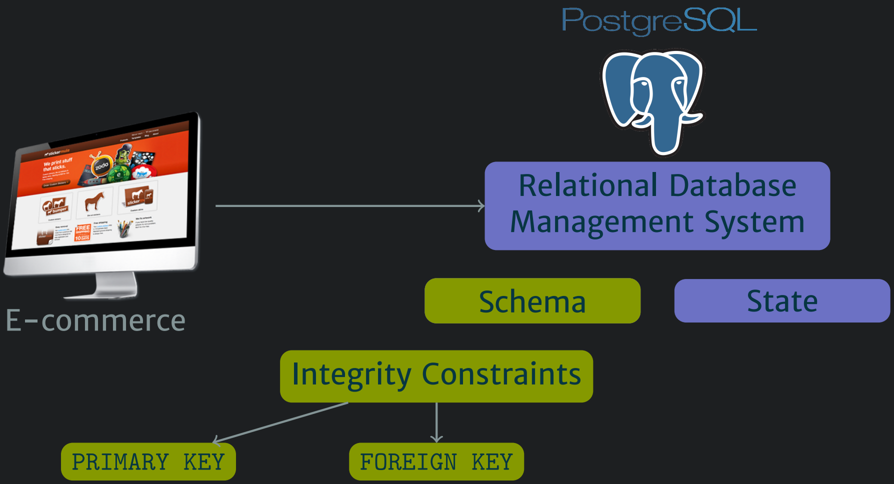
State

Integrity Constraints

# Database Schemas



# Database Schemas



# Database Schemas

E-commerce

PostgreSQL

Relational Database Management System

Schema

State

Integrity Constraints

`PRIMARY KEY`

`FOREIGN KEY`

Arbitrary `CHECK`

# Database Schemas

PostgreSQL

E-commerce

Relational Database Management System

Schema

State

Relational Components

# Database Schemas

PostgreSQL

E-commerce

Relational Database Management System

Schema

State

Relational Components

Tables

# Database Schemas

E-commerce

PostgreSQL

Relational Database Management System

Schema

State

Relational Components

Tables

Rows

# Database Schemas

PostgreSQL

E-commerce

Relational Database Management System

Schema

State

Relational Components

Tables

Rows

Columns

# Database Testing

The Data Warehouse Institute reports that North American organizations experience a **$611 billion annual loss** due to poor data quality

# Database Testing

The Data Warehouse Institute reports that North American organizations experience a **$611 billion annual loss** due to poor data quality

Scott W. Ambler argues that the **“virtual absence” of database testing** — the validation of the contents, schema, and functionality of the database — is the primary cause of this loss

# Database Testing

The Data Warehouse Institute reports that North American organizations experience a **$611 billion annual loss** due to poor data quality

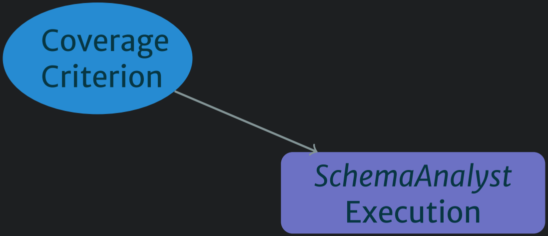
Scott W. Ambler argues that the **“virtual absence” of database testing** — the validation of the contents, schema, and functionality of the database — is the primary cause of this loss

Past papers presented *SchemaAnalyst*, a search-based system for **testing** the complex integrity constraints in **relational schemas**

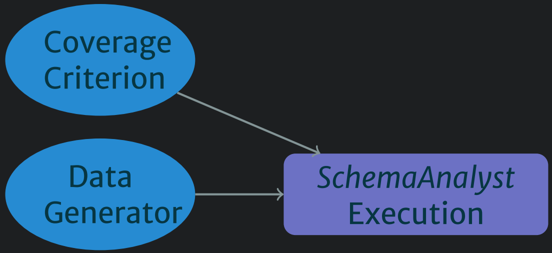
# Method of Approach

*SchemaAnalyst*
Execution

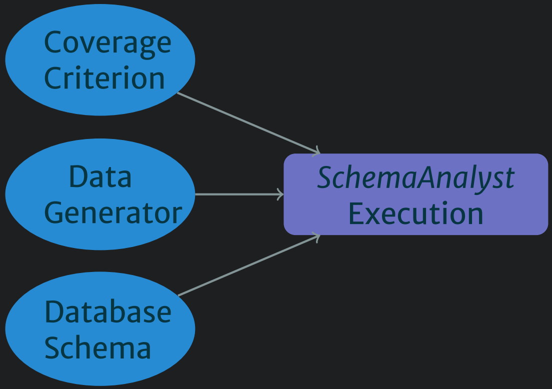
# Method of Approach



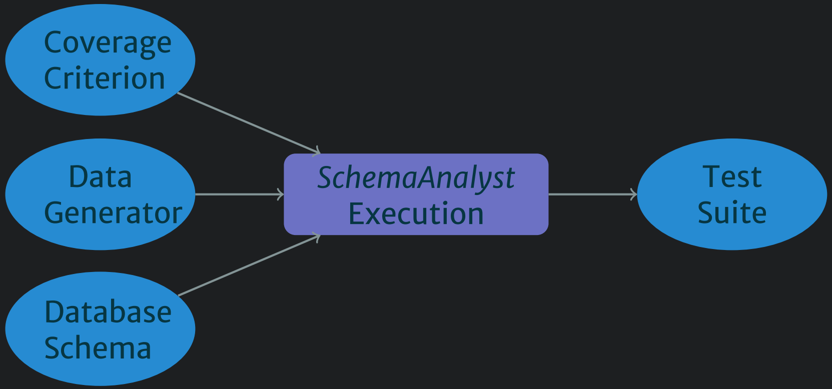
# Method of Approach



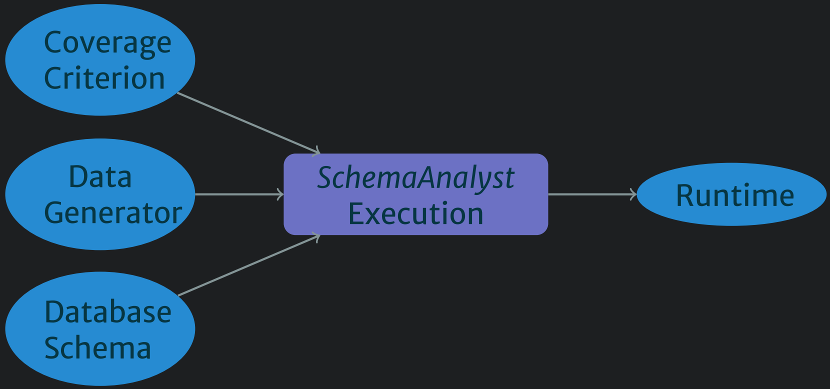
# Method of Approach



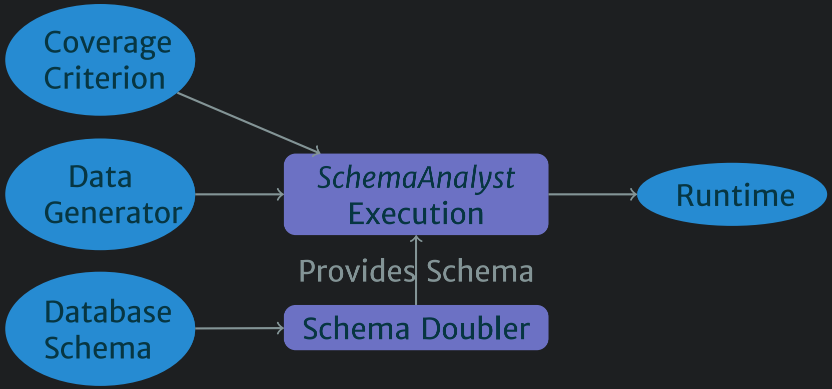
# Method of Approach



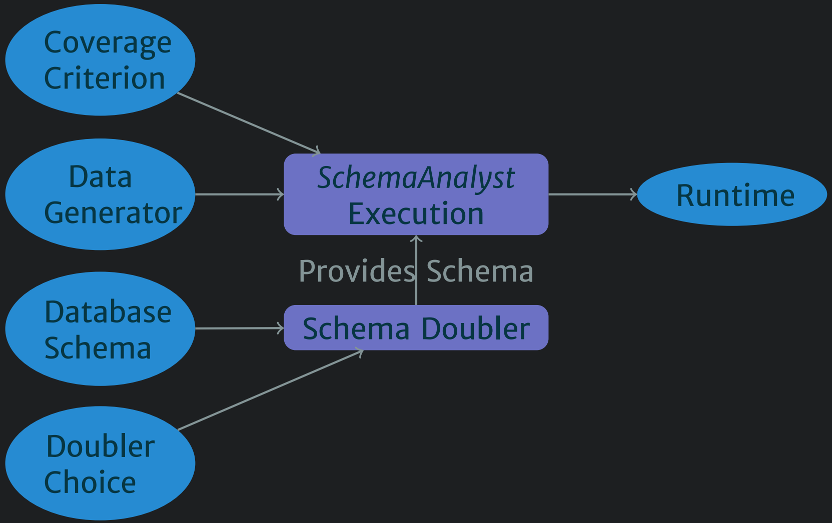
# Method of Approach



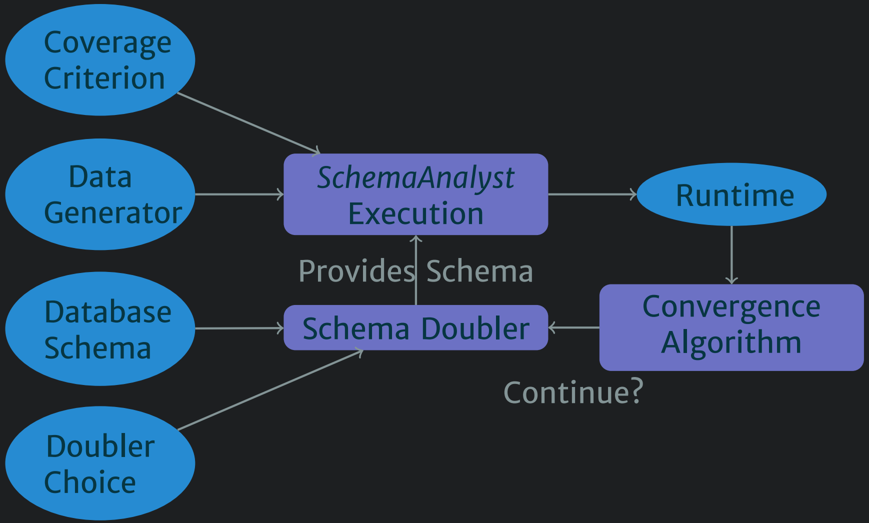
# Method of Approach



# Method of Approach



# Method of Approach

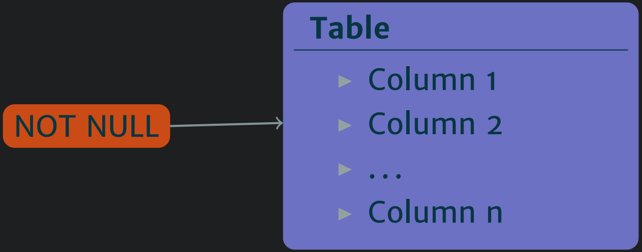


# Doubling Schemas

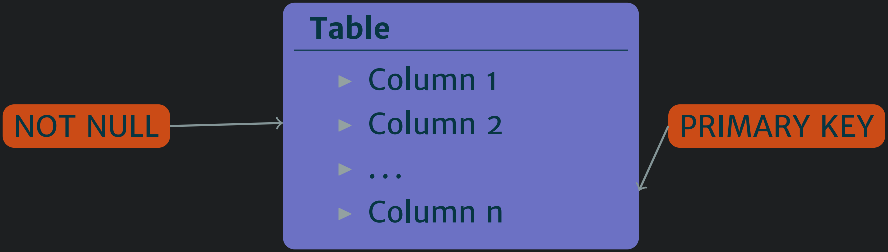
**Table**

- Column 1
- Column 2
- ...
- Column n

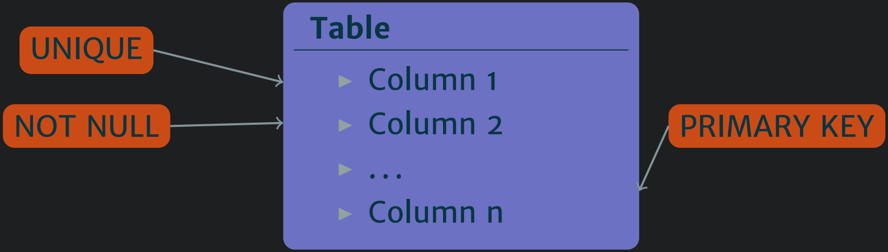
# Doubling Schemas



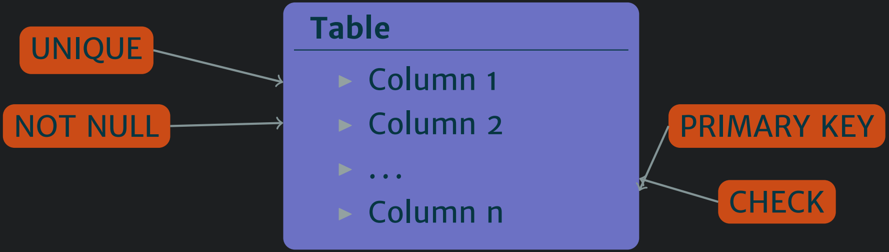
# Doubling Schemas



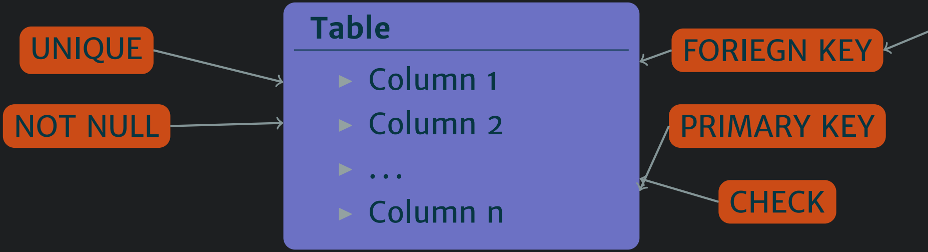
# Doubling Schemas



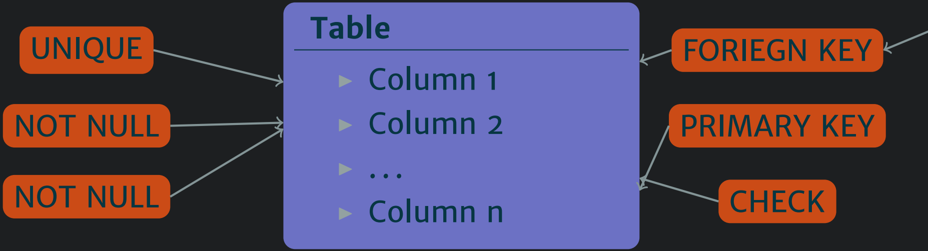
# Doubling Schemas



# Doubling Schemas



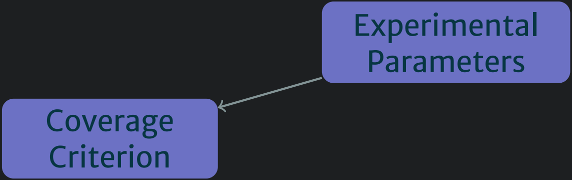
# Doubling Schemas



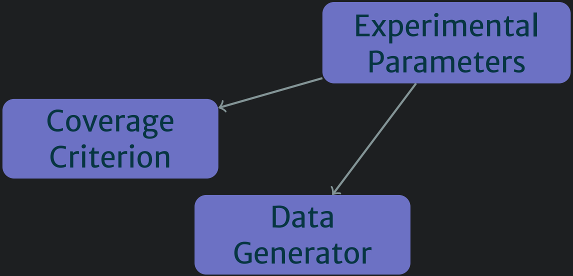
# Experiments

Experimental Parameters

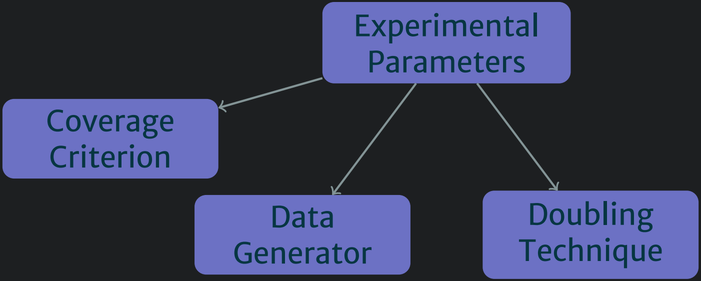
# Experiments



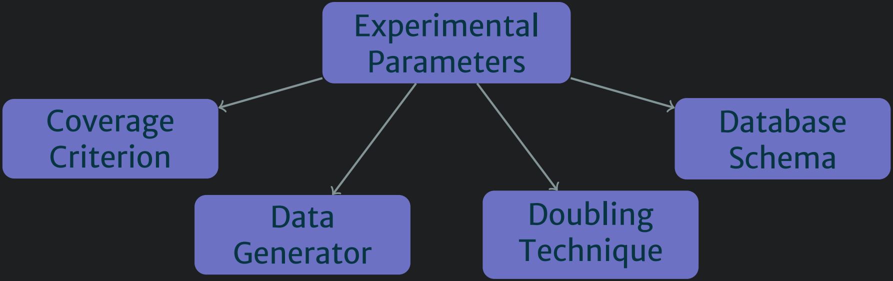
# Experiments



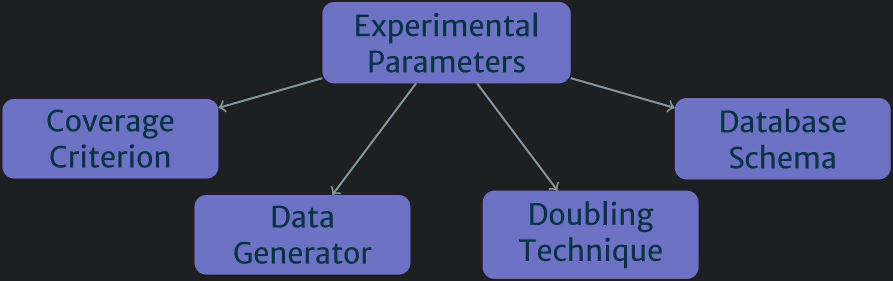
# Experiments



# Experiments



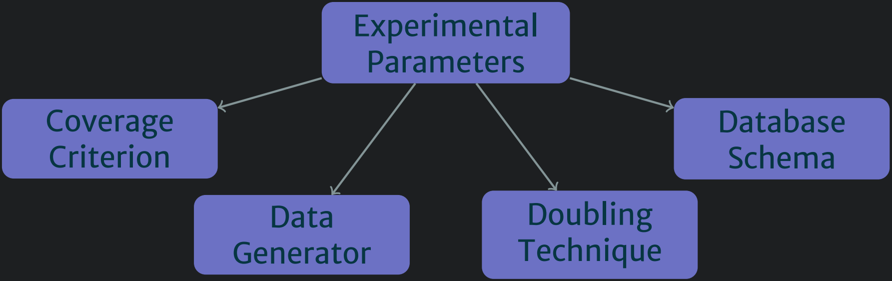
# Experiments




Over 2,000 unique combinations of parameters!

# Experiments

Experiments ran on HPC cluster with 3,440 cores




Over 2,000 unique combinations of parameters!

# Relational Schemas

| Schema | Tables | Columns | Constraints |
|---|---|---|---|
| BioSQL | 28 | 129 | 186 |
| Cloc | 2 | 10 | 0 |
| iTrust | 42 | 309 | 134 |
| JWhoisServer | 6 | 49 | 50 |
| NistWeather | 2 | 9 | 13 |
| NistXTS7 | 1 | 3 | 3 |
| NistXTS749 | 1 | 3 | 3 |
| RiskIt | 13 | 57 | 36 |
| UnixUsage | 8 | 32 | 24 |

# Relational Schemas

| Schema | Tables | Columns | Constraints |
|---|---|---|---|
| BioSQL | 28 | 129 | 186 |
| Cloc | 2 | 10 | 0 |
| iTrust | 42 | 309 | 134 |
| JWhoisServer | 6 | 49 | 50 |
| NistWeather | 2 | 9 | 13 |
| NistXTS7 | 1 | 3 | 3 |
| NistXTS749 | 1 | 3 | 3 |
| RiskIt | 13 | 57 | 36 |
| UnixUsage | 8 | 32 | 24 |

# Relational Schemas

| Schema | Tables | Columns | Constraints |
|---|---|---|---|
| BioSQL | 28 | 129 | 186 |
| Cloc | 2 | 10 | 0 |
| iTrust | 42 | 309 | 134 |
| JWhoisServer | 6 | 49 | 50 |
| NistWeather | 2 | 9 | 13 |
| NistXTS7 | 1 | 3 | 3 |
| NistXTS749 | 1 | 3 | 3 |
| RiskIt | 13 | 57 | 36 |
| UnixUsage | 8 | 32 | 24 |

# Relational Schemas

| Schema | Tables | Columns | Constraints |
|---|---|---|---|
| BioSQL | 28 | 129 | 186 |
| Cloc | 2 | 10 | 0 |
| iTrust | 42 | 309 | 134 |
| JWhoisServer | 6 | 49 | 50 |
| NistWeather | 2 | 9 | 13 |
| NistXTS7 | 1 | 3 | 3 |
| NistXTS749 | 1 | 3 | 3 |
| RiskIt | 13 | 57 | 36 |
| UnixUsage | 8 | 32 | 24 |

# Empirical Results

**Doubled**

- UNIQUEs
- NOT NULLs
- CHECKs

# Empirical Results

**Doubled**

- UNIQUEs
- NOT NULLs
- CHECKs

699 Experiments

# Empirical Results

**Doubled**

- UNIQUEs
- NOT NULLs
- CHECKs

699 Experiments

8% Stopped

# Empirical Results

**Doubled**

- UNIQUEs
- NOT NULLs
- CHECKs

699 Experiments

8% Stopped

20% $O(1)$ or $O(\log)$

# Empirical Results

**Doubled**

- UNIQUEs
- NOT NULLs
- CHECKs

699 Experiments

8% Stopped

20% $O(1)$ or $O(\log)$

72% $O(n)$ or $O(n \log n)$

# Empirical Results

**Doubled**

- UNIQUEs
- NOT NULLs
- CHECKs

699 Experiments

8% Stopped

20% $O(1)$ or $O(\log)$

72% $O(n)$ or $O(n \log n)$

*SchemaAnalyst* $\in O(n)$ for constraints studied

# Empirical Results

**Doubled**

- Tables

# Empirical Results

467 Experiments

**Doubled**

- Tables

# Empirical Results

467 Experiments

**Doubled**

- Tables

56% Stopped

# Empirical Results

| | 467 Experiments |
|---|---|
| **Doubled** | 56% Stopped |
| ▸ Tables | 72 $O(n^2)$ |

# Empirical Results

**Doubled**

- Tables

467 Experiments

56% Stopped

72 $O(n^2)$

10 $O(n^3)$

# Empirical Results

**Doubled**

- Tables

467 Experiments

56% Stopped

72 $O(n^2)$

10 $O(n^3)$

*SchemaAnalyst* $\in O(n^3)$ or worse for tables

# Empirical Results

**Doubled**

- Columns

# Empirical Results

467 Experiments

**Doubled**

- Columns

# Empirical Results

467 Experiments

**Doubled**

- Columns

203 Stopped

# Empirical Results

| | 467 Experiments |
|---|---|
| **Doubled** | 203 Stopped |
| ▸ Columns | 208 $O(n)$ or $O(n \log n)$ |

# Empirical Results

**Doubled**

- Columns

467 Experiments

203 Stopped

208 $O(n)$ or $O(n \log n)$

28 $O(n^2)$ and 2 $O(n^3)$

# Empirical Results

467 Experiments

**Doubled**

- Columns

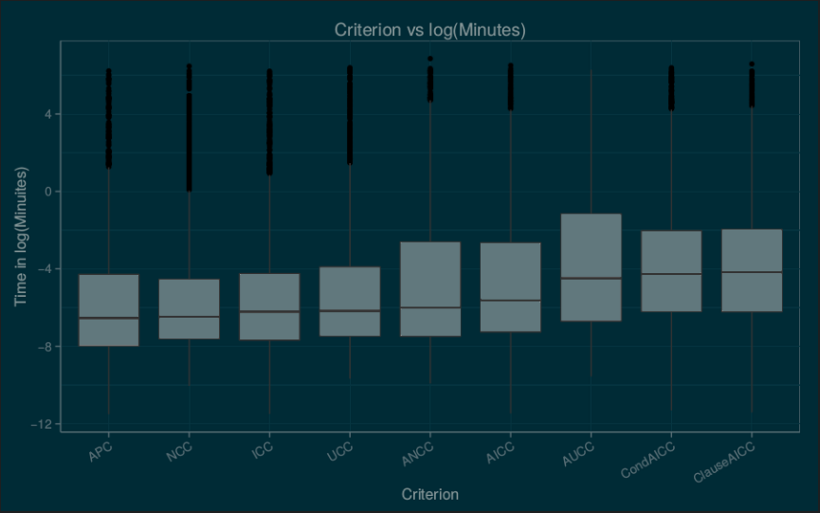
203 Stopped

208 $O(n)$ or $O(n \log n)$

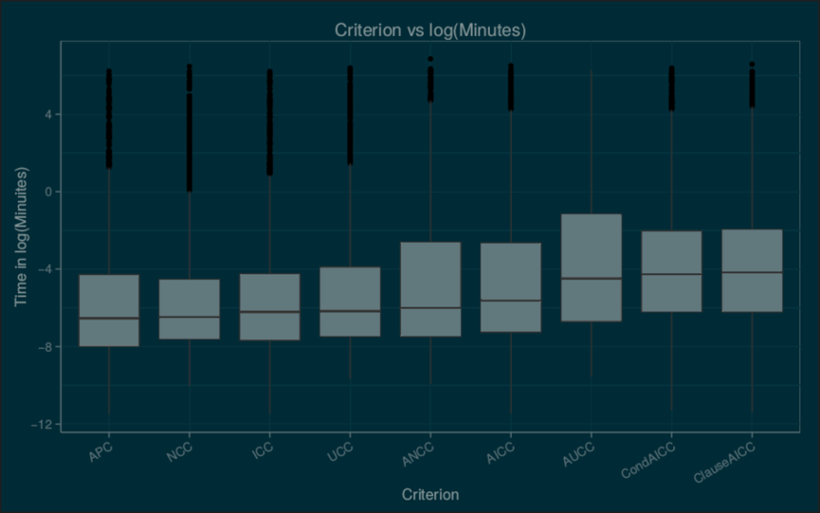
28 $O(n^2)$ and 2 $O(n^3)$

*SchemaAnalyst* $\in O(n^3)$ or worse for columns

# Adequacy Criteria

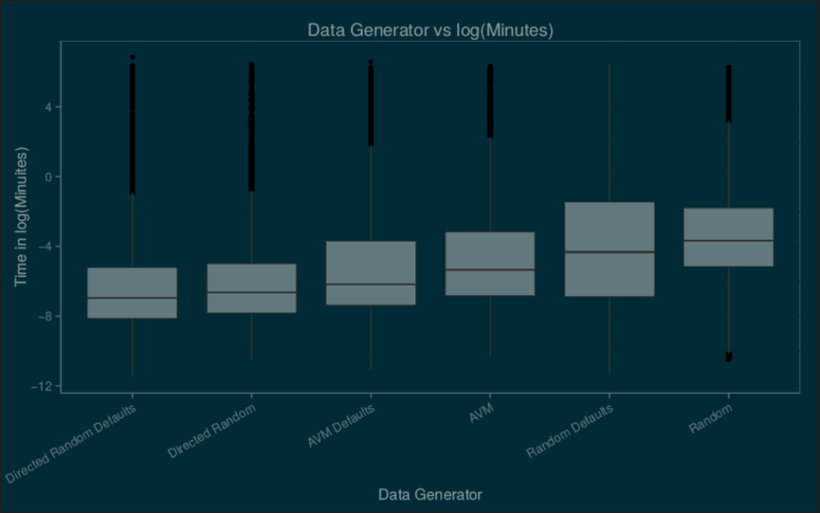


# Adequacy Criteria

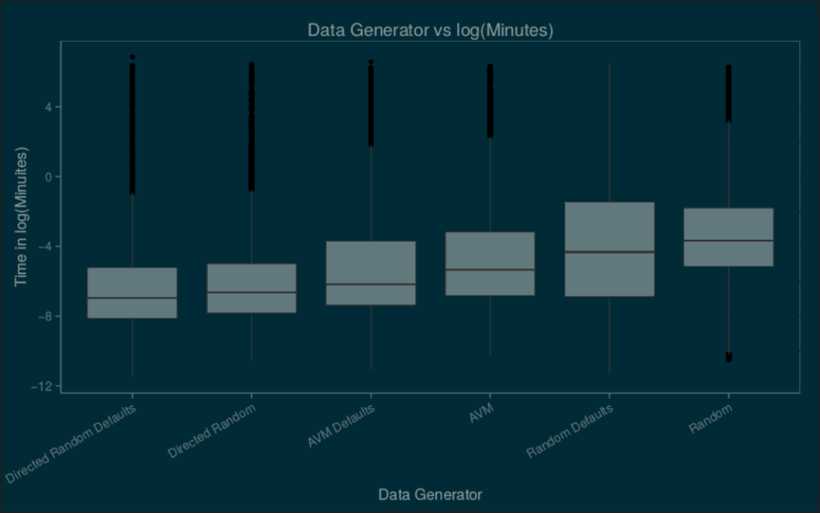



More effective criteria require additional runtime

# Data Generator



# Data Generator




More effective generators can also be more efficient

# Key Contributions

Search-based test data generation is often highly effective, but worst-case time complexity unknown

# Key Contributions

Search-based test data generation is often highly effective, but worst-case time complexity unknown

A technique for automated doubling experiments

# Key Contributions

Search-based test data generation is often highly effective, but worst-case time complexity unknown

A technique for automated doubling experiments

Emprical suggestions for worst-case time complexity

# Key Contributions

Search-based test data generation is often highly effective, but worst-case time complexity unknown

A technique for automated doubling experiments

Emprical suggestions for worst-case time complexity

Tradeoffs in search-based test data generation

# Key Contributions

Search-based test data generation is often highly effective, but worst-case time complexity unknown

A technique for automated doubling experiments

Emprical suggestions for worst-case time complexity

Tradeoffs in search-based test data generation

https://github.com/kinneerc/ExpOse